THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Learning Adaptive Grasping From Human Demonstrations

Citation for published version:
Wang, S, Hu, W, Sun, L, Wang, X & Li, Z 2022, 'Learning Adaptive Grasping From Human Demonstrations', *IEEE/ASME Transactions on Mechatronics*. https://doi.org/10.1109/TMECH.2021.3132465

Digital Object Identifier (DOI):
10.1109/TMECH.2021.3132465

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
IEEE/ASME Transactions on Mechatronics

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 07. Apr. 2022

Learning Adaptive Grasping From Human Demonstrations

Shuaijun Wang, Wenbin Hu, Lining Sun, Xin Wang, Zhibin Li

***Abstract*—This work studied a learning-based approach to learn grasping policies from teleoperated human demonstrations which can achieve adaptive grasping using three different neural network (NN) structures. To transfer human grasping skills effectively, we used multi-sensing state within a sliding time window to learn the state-action mapping. By teleoperating an anthropomorphic robotic hand using human hand tracking, we collected training datasets from representative grasping of various objects, which were used to train grasping policies with three proposed NN structures. The learned policies can grasp objects with varying sizes, shapes, and stiffness. We benchmarked the grasping performance of all policies, and experimental validations showed significant advantages of using the sequential history states, compared to the instantaneous feedback. Based on the benchmark, we further validated the best NN structure to conduct extensive experiments of grasping hundreds of unseen objects with adaptive motions and grasping forces.**

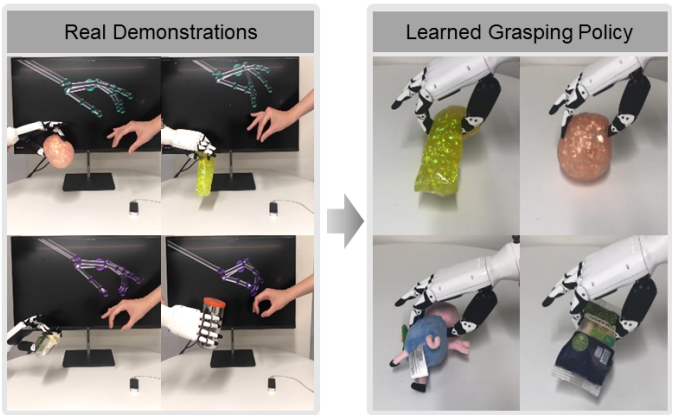


Fig. 1: Learning from human demonstrations using small real-world data to transfer skills for robotic grasping of various objects.

I. Introduction

With limited prior-knowledge of the objects, humans can grasp objects with diverse shapes and stiffness and adapt grasping forces during the interaction. In contrast, for robotic grasping, despite the significant progress in object recognition [1] and grasp synthesis [2] [3] [4], generating adaptive grasping forces to various objects remains an open problem. Most object grasping uses manually designed rules that apply constant joint torques or fix a threshold of motor current, which is practical for most rigid objects but not suitable or adaptive for various deformable and fragile items, and therefore such simple control policies limit the performance, and can potentially damage the objects.

Robotic grasping is an integrated task of object recognition, motion planning and reactive grasp control. Visual information can be used for selecting contact points and pre-grasp poses, but unknown physical properties such as stiffness cannot be inferred from vision. Therefore, during physical interactions, proprioceptive, force, and/or tactile information is needed to generate appropriate adaptive grasping motions and forces, which are crucial for handling a large variety of daily objects with different stiffness.

Conventionally, separate controllers are designed for objects with different physical properties [5], [6], which requires prior-knowledge of target objects, and lacks the adaptability to various object properties. The scope of this paper is to study effective learning to extract policies based on limited real demonstration data, and to achieve local feedback control policies of robotic fingers that can adapt grasping forces to various daily objects.

Learning based methods such as deep reinforcement learning (DRL) have promising performances in dexterous manipulation. However, learning from scratch without demonstrations requires large training data and long training hours [7], which is not a problem in simulation but problematic with real hardware in the loop. Moreover, the unpredictable emergent behaviours can be unnatural or unsafe on real robots [8]. Unlike the traditional problems that can be modeled by rigid body dynamics, for complex physical interactions such as grasping soft deformable objects, there is no high-fidelity and yet computationally efficient simulations to support the trial-error approach of learning in simulation that can be deployed on real systems directly.

Real experimental data is very scarce for grasping various assorted daily objects that are soft, irregular-shaped, deformable, or rigid. To attain sample-efficiency in the cases where only real experimental data are available, an effective approach for the control design is to learn from demonstrations and capture the policies of human grasping skills. Methods for providing human demonstrations include motion capture [9], customized tools [10] [11] or human signals detection device [12] [13]. With appropriate learning frameworks, the real grasping data can be used to train deep neural networks to map multi-modality sensory feedback to the control actions.

Compared to the aforementioned human-robot motion transfer, gesture recognition requires only one depth camera to map human-robot motions, alleviating the complexity in experiment setup and dependency on extra devices. In this paper, to collect demonstration data, we used the vision-based human hand tracking to teleoperate a robotic hand to grasp objects

Shuaijun Wang, Lining Sun are with Harbin Institute of Technology, China. Xin Wang is with Shenzhen Academy of Aerospace Technology, China. Wenbin Hu, Zhibin Li are with School of Informatics, Edinburgh Center of Robotics, University of Edinburgh, UK. Corresponding author: Lining Sun.

with distinct physical properties, as shown in Fig. 1. The proprioceptive data from the demonstrations are recorded and used to extract the underlying human state-action mapping.

The proposed learned-based grasping can produce adaptive grasping forces solely based on the measured proprioceptive data (joint positions and forces), which can be integrated with many existing motion planning algorithms that generate pre-grasp poses, or used for prosthetic hand control. Given guided hand poses, the learned controller can produce fine motor skills with adaptive grasp motions. In contrast to the analytical approaches, learning-based controllers encode the grasping policies into the weights and biases of neural networks, which are efficient to store and easy to update when new policies are learned by augmenting better demonstrations.

To compare the effectiveness of learning from demonstrations, we studied different combinations of the sensory data feedback and different neural network designs. Specifically, spatiotemporal data with history information and instantaneous data without history information are compared and evaluated. We realized effective learning of reactive motor control of the anthropomorphic robotic fingers from few representative demonstrations, with adaptive grasping motions and forces.

The contributions of this paper are as follows:

- A framework of learning adaptive robotic grasping from human demonstrations using an anthropomorphic robotic hand;
- Study and evaluation of the effectiveness of different state feedback for learning the state-action mapping;
- Comparison study of three neural network structures with and without the history input, and their performance validation by grasping objects with various shapes and stiffness.

The paper is organized as follows. Section II introduces the related works. Section III presents the teleoperation system and the collection of human grasping demonstrations. Section IV presents the methodology, including the state combination analysis, controller design, policy learning and the policy evaluation. The experimental results are presented in Section V. The discussion is given in Section VI. Finally, we draw the conclusion and discuss future work in Section VII.

II. Related work

A. Analytic methods

Conventional grasping controllers are designed using analytic models based on the feedback of actuator torques and positions [14]–[17], but subject to limited adaptive ability, especially in grasping objects with various physical properties.

Pfanne et al. proposed an object-level impedance controller for dexterous in-hand manipulation capable of handling dynamic changes in the grasp configuration [14]. The proposed algorithm in [18] switches between force and position control according to the external force. Romano et al. introduced a framework which divided the grasping process into discrete phases based on the tactile information [19]. Most of this paradigm of solutions are based on human ingenuity and handcraft of control rules [20].

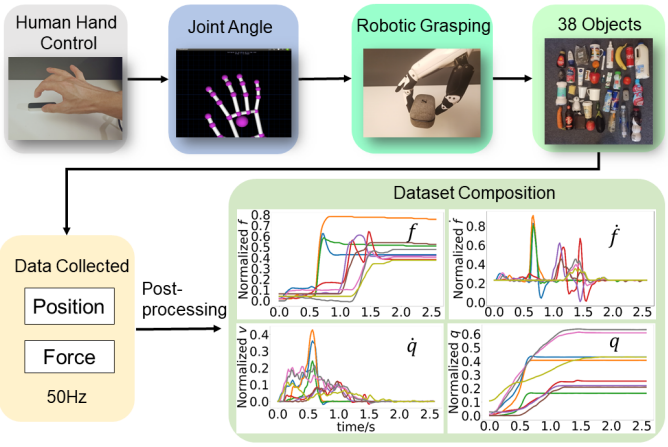


Fig. 2: Dataset collection for teleoperated learning from demonstrations.

B. Learning-based methods

Many recent work on autonomous grasping used learning-based approaches [21]–[24]. Lee et al. used multimodal sensory fusion, including visual, kinesthetic and ontological information, to achieve decision making and continuous control based on self-supervised learning [25]. Shahid et al. [26] leveraged deep reinforcement learning to train a unified policy for reaching, grasping, and lifting the objects in the simulation. The work in [27] used spatio-temporal information and an long short-term memory (LSTM) network to learn the classification of object materials and grasping phases. The high-dimensional saptio-temporal human grasping data can be embedded into a lower-dimensional space for modeling and recognition of grasping actions [28], which demonstrated the effectiveness of saptio-temporal data in learning grasping policies. Also, the history data is effective in slippage detection [27], [29].

C. Learning from demonstration for robotic grasping

Learning the robotic grasping from human demonstrations has been widely researched recently [30], [31]. Misimi et al. used the support vector regression (SVR) to formulate the grasping controller with both vision and tactile feedback, in order to grasp the compliant food objects [32]. In [33], the researchers used sequential grasping data to achieve grasp recognition, which shows the potential of history proprioceptive data to realize autonomous grasping. Most of the prior works using human demonstrations focused on the grasp planning that generates the pre-grasp pose and contact points for the target object [10], [32], [34]. While little attention has been paid to the finger control and the grasp execution. In this paper, we utilize the history data of multi-sensing feedback in the framework of learning from human demonstrations, to directly learn the policy of continuous, adaptive grasping of a wide range of objects.

III. Collection of human demonstration data

This section presents the tele-operation system and the collection of human demonstration data, and the composition of the dataset that used for training.

1) Teleoperation and data collection: We use the leap motion hand tracking device to detect real-time human fingers motion, and teleoperate the robotic hand via kinematics mapping from human fingers to robotic actuators, as shown in Fig. 2. Then the robot proprioception data including actuators forces, positions and their first order derivatives are recorded as the training data. In this work, the thumb, index and middle fingers are used for grasping a single object.

At the beginning of each demonstration, given a randomly placed object, the robot hand is held and placed at a proper pre-grasp pose by the human operator. During the grasp, the pose of the robotic hand remains fixed. Then, the robotic fingers are teleoperated by the demonstrator to grasp the object, with adaptation to the object's physical properties. Note that the human hand is merely providing the grasping motion, without grasping any real object.

The teleoperation approach of providing demonstrations can mitigate the discrepancies between robot hand and human hand, because the operator can learn to adapt motor skills such that the reflected skills at the robot side are feasible and suitable for the robot itself, so as to ensure successful grasping. Though the robot hand has less degrees of freedom compared to human hand, our trials show that the robotic three-finger grasping motion is very similar to that of humans, since the robot hand has a similar size and morphology to the human hand.

2) Composition of dataset: The training data consists of the forces and positions of the linear actuators that drive the fingers during the grasping motion. 38 objects with variations in shape, size and stiffness are used to provide the grasp demonstrations. To fully take advantage of the data, the K-fold cross-validation method is applied in the training. We randomly choose 5 objects as the validation dataset and the rest is the training dataset.

We define q and f as the measurements of positions and forces of the linear actuators, and hereby their computed derivatives are $\dot{q}$, $\dot{f}$ respectively. The linear actuator drives the intermediate linkage mechanism, which then enables the revolute finger joint to rotate. Note that the force sensor is placed outside of the drive chain so the measured forces are directly applied to the finger joints via the linkage mechanism.

During the online grasping, the q and f are recorded at 50Hz and post-processed by lowpass filters (first-order with cutoff frequency of 10 Hz).

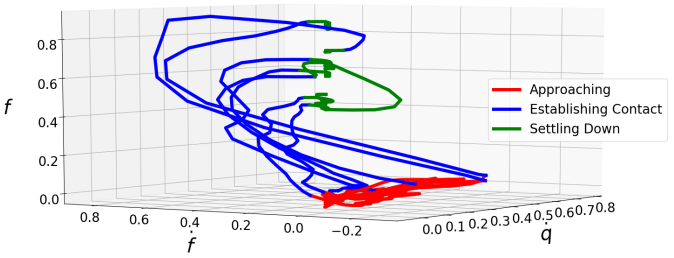
IV. Method

A. Description of proprioceptive policies

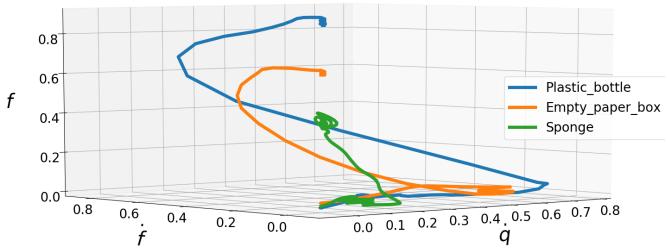
The policy of adaptive grasping which maps the robot proprioception to the control signals, can be represented as

$$\dot{q}_{\mathrm{d}} = \pi(s_t), \tag{1}$$

where $\dot{q}_{\mathrm{d}}$ is the desired velocities of the finger linear actuators, which are monotonic with the finger joint velocities. s_t denotes the vector of the state feedback at t time (see more in Section IV-C); π denotes the policy that maps the state feedback to the desired actions, which is represented by a neural network trained from human demonstration data.



(a) Grasping phases: different colors represent different phases.



(b) Grasping of objects of different stiffness.

Fig. 3: Normalized feedback states from the index finger during the grasping of three representative objects.

B. Framework of learning from grasping demonstrations

Fig. 4 demonstrates the algorithm framework which consists of three modules: data generation, offline training, and online grasping. The training dataset is obtained from human demonstrations in the data generation module. Given the training data, the grasping skills from human demonstrations are transferred to the NN-based controllers via supervised learning in the offline training module.

We propose and evaluate three controllers with different structures (see more in Section IV-D). Once trained, the learned NN-based policies are used in the online grasping module as a feedback controller, where the measured robot proprioceptive data are post-processed and fed as input. As shown in Fig. 4, the outputs are the desired velocity commands for the finger actuators.

C. Analysis of grasping data and state combination selection

In this section, we analyse the characteristics of robot proprioceptive data during grasping, in order to provide theoretical support for the selection of effective input states combination for learning the policy.

1) Data analysis of the grasping process: During the grasping, the state vector $[\dot{q},\ f,\ \dot{f}]$ can be used to distinguish grasping phases. The actuator position q is less indicative, because it can not demonstrate the phase of establishing contact, and the equilibrium of q depends on the shape/size of the object which is unnecessary for our controller. When a grasp reaches the equilibrium, $\dot{q}$ and $\dot{f}$ converge to zero, and f converges to a settled value. $\dot{q}, \dot{f}$ can encode the information of object stiffness during the early stage of contact. f indicates the grasping force, and $f, \dot{f}$ can reflect the contact transitions. Therefore, the tuple of state vector $[\dot{q},\ f,\ \dot{f}]$ is used for the policy learning.

Fig. 3 shows representative trajectories of $[\dot{q},\ f,\ \dot{f}]$ from the index finger during the grasping of different objects. In

TABLE I: Clustering analysis on different feedback states.

NO. of clusters	$f, \dot{f}$	$\dot{q}, f$	$\dot{q}, \dot{f}$	$\dot{q}, f, \dot{f}$	$q, \dot{q}, f, \dot{f}$
10	0.0058	0.0060	0.0049	**0.0061**	0.0051
12	0.0058	0.0086	0.0112	**0.0115**	0.0051

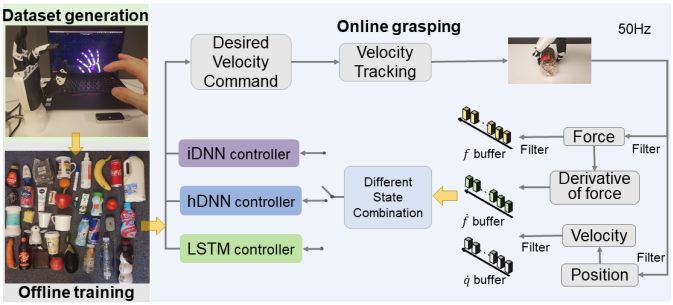


Fig. 4: The proposed framework include three parts: data generation module, offline training module and online grasping module.

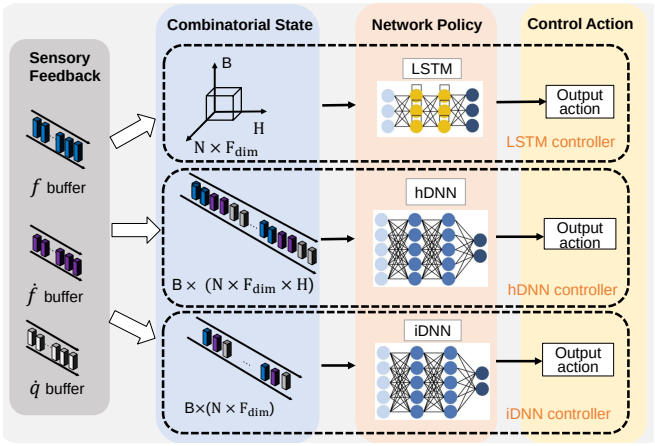


Fig. 5: Three network structures and their feedback states.

this state space, grasping different objects shows different trajectories, where the process can be categorized into 3 sequential phases as: approaching, establishing contact, and settling down.

- *Approaching:* $\dot{f}$, f are around 0, and $\dot{q}$ becomes non-zero due to the movement.
- *Establishing contact:* $\dot{f}$, $\dot{q}$, f are evolving during this transition. $\dot{f}$ and $\dot{q}$ will rise from 0, reach peaks and then drop to 0. f will rise and reach a constant. Objects with different sizes and stiffness will result in distinct and different trajectories in the state space.
- *Settling down:* f maintains at a constant value, while $\dot{f}$, $\dot{q}$ settles around 0.

2) Clustering analysis of feedback states: To select the most effective combination of input states that can differentiate different physical interaction phases with different objects, the complete training dataset is partitioned into a number of clusters (10 and 12 are used here based on the scale of the dataset) by the unsupervised clustering method *K-means* [35]. Every data-point for clustering is the temporal state feedback within a fixed time-window $s_{t-H:t}$, here H refers to the size of time window. The clustering results are evaluated by *Dunn index* [36], i.e. a larger number indicates the more distinguishable clusters. Table I shows that the state vector $s_t = [\dot{q}, f, \dot{f}]$ has the highest Dunn indexes and captures at least 10 to 12 most distinct phases in temporal sensory measurements.

3) Selection of state input: Based on the grasping data analysis and clustering analysis above, the state combination $s_t = [\dot{q}, f, \dot{f}]$ can differentiate object properties and characterize the grasping phases. To further evaluate the effectiveness of history data, we designed two types of state input for policy learning: (1) the instantaneous state vector s_t at the current timestep, and (2) the temporal state tuple $s_{t-H:t}$ using history data within a fixed time window.

D. Design of grasping controllers

As for the effective skill transfer, we used supervised learning, which is computationally efficient to train the grasping policy π directly with the demonstration data. To focus on the evaluation of history information, and to alleviate any influence to the results introduced by the network structures, we used the simplest network structure – fully connected neural network – as the structure of iDNN (DNN with instantaneous information) using s_t, and hDNN (DNN with history information) using $s_{t-H:t}$. LSTM network is widely used in processing sequential data. Hence we also designed an LSTM-based controller using history input $s_{t-H:t}$. The time window used in this work is 0.4s, which can cover the transition phase of contact in most robotic grasping. Moreover, an over-long history will include unneeded information and increase computation, while a too-short history is not enough to distinguish different grasping phases. Empirically, we choose this parameter based on the empirical knowledge of the average contact phase during most tasks. The detailed structures and state inputs of three controllers are shown in Fig. 5.

1) iDNN: The input is the instantaneous state $s_t = [\dot{q}, f, \dot{f}]$, with dimension $I_1 = B \times (N \times F_{dim})$, where B denotes the batch size. N denotes the degree of freedom and F_{dim} denotes the feature dimension. The output is the finger action vector with dimension $N \times 1$. The network has two fully connected hidden layers.

2) hDNN: The input is the temporal state tuple $s_{t-H:t}$ including the *history* state within a time window, with dimension $I_2 = B \times (N \times F_{dim} \times H)$, where H denotes the size of time window. Except the input size, the rest of the network structure is the same with iDNN, with two fully connected hidden layers and one output layer.

3) LSTM: Recurrent neural network (RNN) is applied to construct the LSTM grasping controller. The input state with dimension $I_3 = B \times (N \times F_{dim}) \times H$ is fed into two LSTM layers and one fully connected output layer.

E. Policy learning

With the collected training dataset, the aforementioned controllers are trained via supervised learning. The loss function is defined as mean squared errors between the ground truth and the output of control actions ($\dot{q}_d$) plus L2 regularization:

$$loss = \sum_{i=1}^{n}(y_i - y_i^d)^2/n + \lambda \sum_{i=1}^{k} \omega_i^2, \qquad (2)$$

Fig. 6: Unseen objects with various physical properties for testing.

where y_i is the ground truth of finger actions at timestep i, and y_i^d denote the network's output actions. n is the number of samples. ω_i is the weight of the network, and λ is 0.001.

F. Evaluation of robustness

To evaluate the robustness of successful grasps against perturbations, we define γ as the force metric which is the resultant force generated by all the actuators of the fingers:

$$\gamma = \sqrt{\sum_{i=1}^{n} f_i^2}, \tag{3}$$

where f_i is the measured force of each actuator, n is the number of fingers. In this paper, we only consider the grasping of general daily objects, excluding fragile objects like empty egg shells. Hence as long as the grasping forces do not damage the object, this metric can indicate the resistance against external disturbances. In the following experiments, the force metric γ is used to measure the robustness of each grasp of different objects.

V. Experimental Results and Analysis

A. Hardware setup

An anthropomorphic multi-fingered hand, Inspire Robot Hand, is used for experiments. The hand has six degrees of freedom, two for the thumb and one for each of the remaining fingers. Each degree of freedom is driven by a linear actuator with a one-dimensional force sensor and a position sensor mounted at the output of the drive chain, detecting the real-time force and position of the motor. We have a customized control loop running at 50Hz to update the position references of an internal position-control loop to achieve velocity control.

B. Grasping experiments with unseen objects

TABLE II: Success rates for grasping experiments and effective state combinations for policy learning.

	iDNN	hDNN	LSTM
Success rates	93%	94%	88%
Valid combinations	None	$\dot{q}$, f+$\dot{q}$, $\dot{f}$+$\dot{q}$, f+$\dot{f}$+$\dot{q}$	f, f+$\dot{f}$, f+$\dot{q}$, f+$\dot{f}$+$\dot{q}$

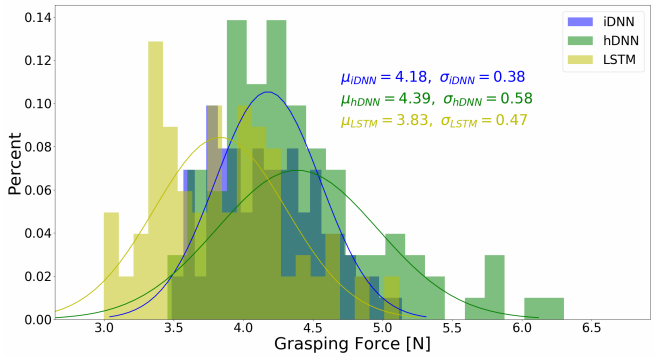
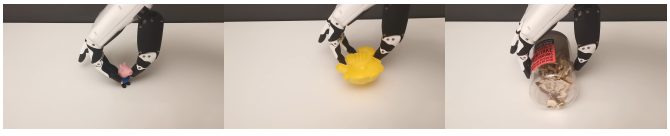


Fig. 7: Distribution of grasping forces from different controllers over the testing objects, including the mean and standard deviation.



(a) Small toy (b) Plastic model (c) Bottle

(d) Soft toy (e) Peeled boiled egg (f) Pepper bottle

Fig. 8: Grasping objects with increasing sizes (top row) and increasing stiffness (bottom row).

To evaluate the robustness and generalization of the learned controllers, we conducted grasping experiments on 100 unseen objects with various sizes, shapes and stiffness, as shown in Fig. 6. The wrist of robotic hand is positioned by an operator, who selects the grasping pose, and the rest of in-hand grasping is executed by the learned controllers. The success rates of three controllers are in Table II, and hDNN controller has the best performance with the success rate of 94%.

Fig. 7 shows the distribution of force metric γ over the grasping experiments of unseen objects with three proposed controllers. Despite of a high success rate, the iDNN controller has a smaller standard deviation of γ values, indicating a smaller range of adaptation of forces to different objects. LSTM controller has higher standard deviation and better adaptability to objects with different stiffness, but has the lowest average γ value, which is less robust against uncertainties. Compared with iDNN and LSTM controllers, the hDNN controller is more versatile – on average, it generates larger grasping forces and also has a wide range of force adaptation.

C. Comparison study between three controllers

This section presents the results from grasping 6 distinct and representative objects to demonstrate the performances of learned controllers, as shown in Fig. 8.

1) Grasping objects with similar stiffness and different sizes: As shown in Fig. 9 (a), hDNN and LSTM controllers generate distinct output actions at different grasping phases, but iDNN controller generates relatively constant output during the grasping. As shown in Fig. 9 (b), compared with iDNN,

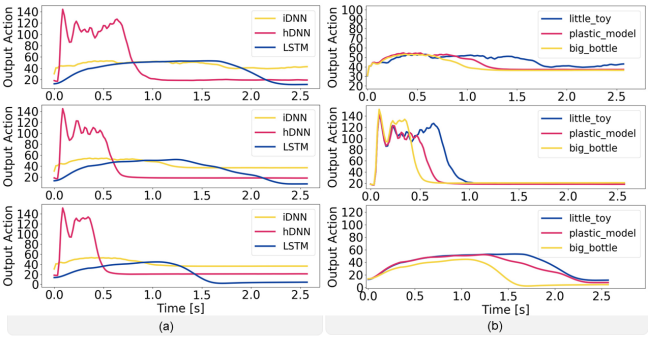


Fig. 9: Comparison of three controllers. (a) Output actions (actuator velocity of index finger) for grasping a small toy, a plastic model, and a bottle respectively (top to bottom); (b) The same data sorted by iDNN, hDNN, LSTM controllers respectively (top to bottom).

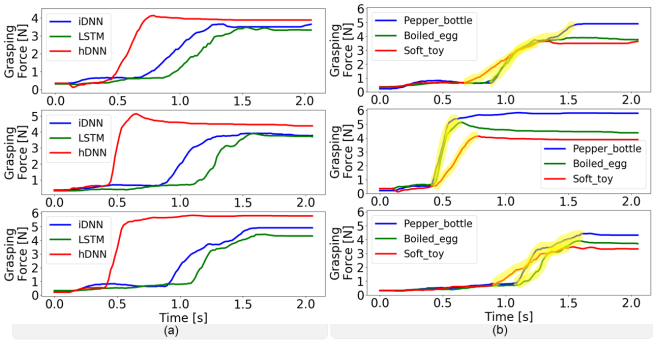


Fig. 10: Comparison of three controllers. (a) Grasping forces during grasping a soft toy, a peeled boiled egg, and a pepper bottle respectively (top to bottom). (b) The same data sorted by iDNN, hDNN, and LSTM controllers respectively (top to bottom), and the transitions are in yellow highlight.

the time-series action curves of hDNN and LSTM are more distinct for different objects shown in Fig. 8(a) - (c), indicating that the history data contributes to the disambiguation of object sizes, and can potentially lead to better adaptability.

2) Grasping objects with similar sizes and different stiffness: For grasping objects shown in Fig. 8(d) - (f), though the finger joint configuration is similar once settled, the dynamic transitions are very different during the contact. With $s_{t-H:t} = [\dot{q},\ f,\ \dot{f}]$ capturing such transitional features within a time window, Fig. 10 (b) shows that hDNN and LSTM controllers have distinct force adaptations, e.g. 3 different stable grasping forces for 3 different objects. The LSTM controller has smaller grasping forces γ than the other two controllers in general as can be seen in Fig. 10 (a), resulting in softer grasps, which is consistent with the statistical results in Fig. 7.

D. Ablation study

To evaluate the effect of each feedback as the controller input, the ablation study has been conducted. Fig. 11 shows the profiles of output actions during grasping the pepper bottle using the three proposed controllers trained with all the possible state combinations. Empirically, the controllers that generate distinct output actions at different grasping stages are regarded as adaptive and reactive, and the corresponding state combinations are effective. As demonstrated in the Fig. 11(a),

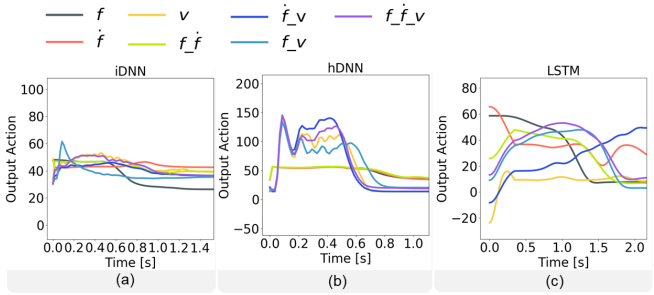


Fig. 11: Output velocity of index finger during grasping the pepper bottle using three controllers trained with different state inputs.

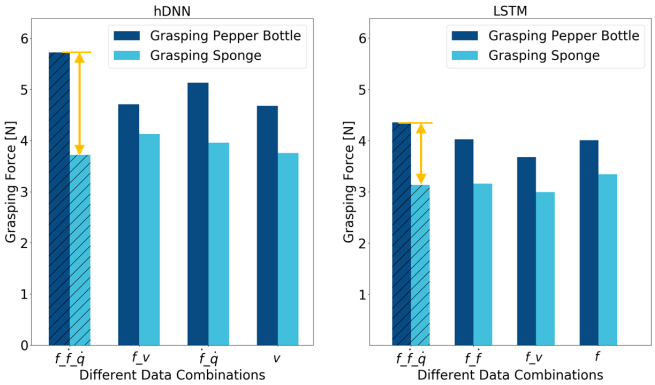


Fig. 12: Settled grasping force using different state feedback from the hDNN (left) and LSTM (right) controller respectively. The combination in the left bar filled with slash has the best performance.

none of the iDNN controllers are adaptive, generating relatively constant actions during the grasping; While in Fig. 11(b) and (c), some state combinations are effective in training adaptive hDNN and LSTM controllers. The effective state combinations for three controllers are listed in Table II.

The iDNN controller using instantaneous feedback s_t without the history data merely generated constant finger actions, which suggests that s_t does not capture sufficient information for encoding human grasp skills. Contrarily, with history information $s_{t-H:t}$, both hDNN and LSTM controllers achieve more human-like grasping, though the effective input combinations vary as shown in Table II.

Further, Fig. 12 compares grasping forces over rigid and deformable objects, suggesting that the combination of multi-sensory data is more effective in learning adaptive grasping. For hDNN and LSTM controllers, the learned policies using the complete state combination $[\dot{q}, f, \dot{f}]$ can generate the most distinct grasping forces for rigid and soft objects.

E. Comparison with the baseline controller

To evaluate the effectiveness and adaptiveness of the learned controllers, we compared them with a pre-programmed baseline controller, which generates constant joint velocities and has a threshold on the grasping force computed as in Eq. (3). Once the grasping force exceeds the threshold, fingers will stop moving and maintain the current joint positions. The force threshold is pre-defined and fixed during the experiments. A paper card is chosen as the target object so that its deformation can be visually observed to evaluate the grasping performance.

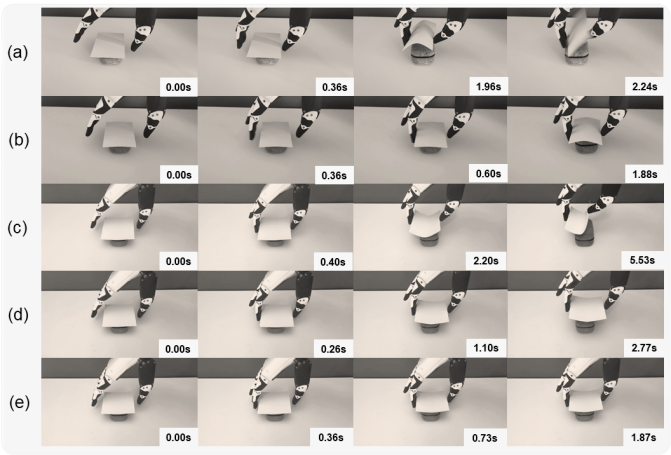


Fig. 13: Comparison by grasping the unseen card: (a-b) Baseline controller with a high and low force thresholds, respectively; (c) iDNN controller; (d) hDNN controller; (e) LSTM controller.

Fig. 13 (a) and (b) show the grasping motion of the baseline controller with a high (4N) and low (2N) force threshold respectively. The difference in the performances shows the importance of a proper force threshold, which requires the prior-knowledge of the object. The iDNN controller generated excessive forces and bent the card, leading to a failure and poor adaptation to low object stiffness during the interaction. In contrast, hDNN and LSTM controllers can hold the card stably without prior-knowledge of the card's physical properties, indicating that they have certain adaptability to the unknown object stiffness and can generate different grasping forces while interacting with different objects. On contrary, the iDNN controllers can only apply constant grasp forces with no self-adaptation, and the baseline controller requires a properly tuned force threshold.

F. Similarities of human and robot policies

The aforementioned results indicate that the history states play an important role in distinguishing the robot grasping phases, and encoding latent information of object shape and stiffness, which enables adaptive grasping of various objects.

Though hDNN and LSTM controllers have comparable performance, the former is better because it has a simpler network structure, larger grasping force and better adaptability to various objects. We constructed the nearest sample neighbours by t-distributed stochastic neighbor embedding (t-SNE) as shown in Fig. 15, using actuator measurements from human tele-operated demonstrations and hDNN-based grasping. The visualisation of large overlapping areas suggest the underlying similarities between the human and learned policies, as well as the effectiveness of history states in representing and extracting the state-action mapping from human grasp policies.

G. Investigation of failure cases

The success of grasping an object depends on the selection of contact points by the user, especially for the positioning the fingertips. Fig. 14 and Fig. 16 demonstrate both the success and failed grasping of representative objects using hDNN controller. Fig. 16(a) and (b) show the failures caused by unbalanced and unstable contact points. Due to the characteristics of point contact, it is also difficult to grasp heavy and slippery objects, as shown in Fig. 16(c).

VI. Discussion

In this paper, we focused on the dexterous grasping and adaptive control of the robotic fingers, which is important while lacking of the prior-knowledge of the object's material and stiffness. With human demonstrations, the proposed hDNN grasping controller is capable of generating adaptive forces to grasp objects with various sizes and stiffness, solely based on the robot proprioception data.

The proposed grasping controller requires the minimum user input: merely a 0-1 activation to start and stop the grasp motion. Therefore the controller can be implemented in many scenarios, e.g. teleoperation system or prosthesis system where the arm motion is controlled by the user and the grasping motion is executed by the proposed controller, which alleviates the operator's mental load from complex grasping control. Also, the grasping controller can be integrated with any off-the-shelf grasp planning algorithms.

VII. Conclusion and Future Work

In this paper, we proposed a learning-based approach of adaptive grasping with an anthropomorphic robotic hand based on few real human demonstrations. We studied different multi-sensing state combinations to encode the state-action mapping of human grasping skills. Three different neural network structures are designed to compare the effectiveness of using the instantaneous and history states in the policy learning. The ablation study and data analysis showed the importance of history data in differentiating grasping phases and generating more robust and adaptive grasping actions. Finally, we extensively tested the learned *hDNN* controller with 100 unseen objects. The experimental results showed that the learned controller was capable of grasping objects with different shapes and stiffness, based on the transferred state-action mapping.

One future extension is to integrate the adaptive grasping controller with grasp planning algorithms which generate suitable pre-grasp poses given vision-guided object semantics. Therefore, more automatic "reach and grasp" motion can be integrated. Furthermore, we will study the usage of more sensory feedback, e.g. tactile and visual information, to improve environmental perception and enable the learning of more intelligent and versatile grasping policies.

Acknowledgment

This work is supported in part by the RSE-NSFC program from Royal Society of Edinburgh, H2020 project Harmony (101017008), and National Natural Science Foundation of China (61911530250).

Fig. 14: Grasping a variety of objects with different physical properties using the trained *hDNN* controller (see accompanying video).

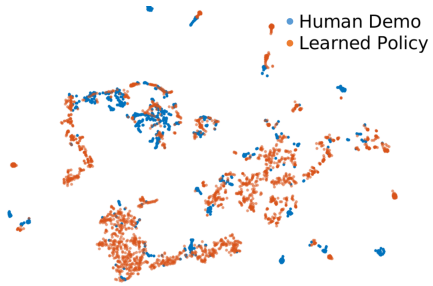


Fig. 15: T-SNE analysis map of human and robotic grasping policies.

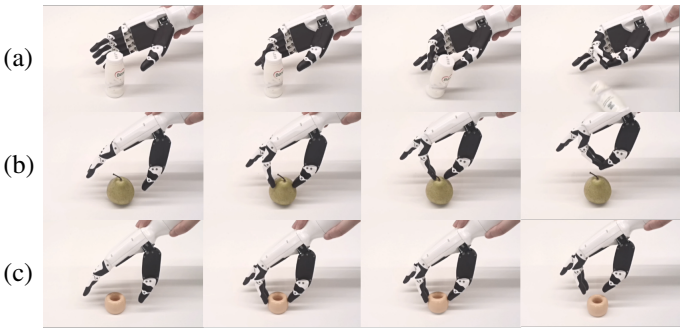


Fig. 16: Representative failures of grasp attempts.

REFERENCES

[1] M. Gou, H. Fang, Z. Zhu, S. Xu, C. Wang, and C. Lu, "RGB matters: Learning 7-dof grasp poses on monocular RGBD images," *CoRR*, vol. abs/2103.02184, 2021.

[2] J. Mahler, J. Liang, S. Niyaz, M. Laskey, R. Doan, X. Liu, J. A. Ojea, and K. Goldberg, "Dex-net 2.0: Deep learning to plan robust grasps with synthetic point clouds and analytic grasp metrics," in *Robotics: Science and Systems (RSS)*, 2017.

[3] D. Kalashnikov, A. Irpan, P. Pastor, J. Ibarz, A. Herzog, E. Jang, D. Quillen, E. Holly, M. Kalakrishnan, V. Vanhoucke, and S. Levine, "Scalable deep reinforcement learning for vision-based robotic manipulation," in *Proceedings of The 2nd Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, vol. 87, 29–31 Oct 2018, pp. 651–673.

[4] D. Morrison, J. Leitner, and P. Corke, "Closing the Loop for Robotic Grasping: A Real-time, Generative Grasp Synthesis Approach." Robotics: Science and Systems Foundation, aug 2018.

[5] M. T. Mason, A. Rodriguez, S. S. Srinivasa, and A. S. Vazquez, "Autonomous manipulation with a general-purpose simple hand," *The International Journal of Robotics Research*, vol. 31, no. 5, pp. 688–703, apr 2012.

[6] T. Yoshikawa, "Multifingered robot hands: Control for grasping and manipulation," *Annual Reviews in Control*, vol. 34, no. 2, pp. 199–208, dec 2010.

[7] O. M. Andrychowicz, B. Baker, M. Chociej, R. Józefowicz, B. McGrew, J. Pachocki, A. Petron, M. Plappert, G. Powell, A. Ray, J. Schneider, S. Sidor, J. Tobin, P. Welinder, L. Weng, and W. Zaremba, "Learning dexterous in-hand manipulation," *The International Journal of Robotics Research*, vol. 39, no. 1, pp. 3–20, 2020.

[8] A. Rajeswaran, V. Kumar, A. Gupta, G. Vezzani, J. Schulman, E. Todorov, and S. Levine, "Learning complex dexterous manipulation with deep reinforcement learning and demonstrations," in *Proceedings of Robotics: Science and Systems*, Pittsburgh, Pennsylvania, June 2018.

[9] F. Ficuciello, A. Migliozzi, G. Laudante, P. Falco, and B. Siciliano, "Vision-based grasp learning of an anthropomorphic hand-arm system in a synergy-based control framework," *Science Robotics*, vol. 4, no. 26, jan 2019.

[10] S. Song, A. Zeng, J. Lee, and T. Funkhouser, "Grasping in the Wild: Learning 6DoF Closed-Loop Grasping from Low-Cost Demonstrations," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4978–4985, jul 2020.

[11] Z. Lu, N. Wang, and C. Yang, "A constrained dmps framework for robot skills learning and generalization from human demonstrations," *IEEE/ASME Transactions on Mechatronics*, pp. 1–1, 2021.

[12] R. Wen, K. Yuan, Q. Wang, S. Heng, and Z. Li, "Force-Guided High-Precision Grasping Control of Fragile and Deformable Objects Using sEMG-Based Force Prediction," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2762–2769, apr 2020.

[13] T. B. Martin, R. O. Ambrose, M. A. Diftler, R. Platt, and M. J. Butzer, "Tactile gloves for autonomous grasping with the NASA/DARPA Robonaut," in *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2004, no. 2. Institute of Electrical and Electronics Engineers Inc., 2004, pp. 1713–1718.

[14] M. Pfanne, M. Chalon, F. Stulp, H. Ritter, and A. Albu-Schäffer, "Object-Level Impedance Control for Dexterous In-Hand Manipulation," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2987–2994, apr 2020.

[15] T. Wimböck, C. Ott, A. Albu-Schäffer, and G. Hirzinger, "Comparison of object-level grasp controllers for dynamic dexterous manipulation," *The International Journal of Robotics Research*, vol. 31, no. 1, pp. 3–23, jan 2012.

[16] M. Pfanne, M. Chalon, F. Stulp, and A. Albu-Schaffer, "Fusing joint measurements and visual features for In-Hand object pose estimation," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3497–3504, oct 2018.

[17] E. Psomopoulou, D. Karashima, Z. Doulgeri, and K. Tahara, "Stable pinching by controlling finger relative orientation of robotic fingers with rolling soft tips," *Robotica*, vol. 36, no. 2, pp. 204–224, feb 2018.

[18] T. Takahashi, T. Tsuboi, T. Kishida, Y. Kawanami, S. Shimizu, M. Iribe, T. Fukushima, and M. Fujita, "Adaptive grasping by multi fingered hand

with tactile sensor based on robust force and position control," in *Proceedings - IEEE International Conference on Robotics and Automation*, 2008, pp. 264–271.

[19] J. M. Romano, K. Hsiao, G. Niemeyer, S. Chitta, and K. J. Kuchenbecker, "Human-inspired robotic grasp control with tactile sensing," *IEEE Transactions on Robotics*, vol. 27, no. 6, pp. 1067–1079, 2011.

[20] M. Kaboli, K. Yao, and G. Cheng, "Tactile-based manipulation of deformable objects with dynamic center of mass," in *IEEE-RAS International Conference on Humanoid Robots*. IEEE Computer Society, dec 2016, pp. 752–757.

[21] J. Bohg, A. Morales, T. Asfour, and D. Kragic, "Data-driven grasp synthesis-A survey," *IEEE Transactions on Robotics*, vol. 30, no. 2, pp. 289–309, 2014.

[22] H. Merzic, M. Bogdanovic, D. Kappler, L. Righetti, and J. Bohg, "Leveraging contact forces for learning to grasp," in *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2019-May, 2019.

[23] W. Yuan, S. Dong, and E. H. Adelson, "Gelsight: High-resolution robot tactile sensors for estimating geometry and force," *Sensors*, vol. 17, no. 12, 2017.

[24] M. Kopicki, R. Detry, M. Adjigble, R. Stolkin, A. Leonardis, and J. L. Wyatt, "One-shot learning and generation of dexterous grasps for novel objects," *The International Journal of Robotics Research*, vol. 35, no. 8, pp. 959–976, 2016.

[25] M. A. Lee, Y. Zhu, K. Srinivasan, P. Shah, S. Savarese, L. Fei-Fei, A. Garg, and J. Bohg, "Making sense of vision and touch: Self-supervised learning of multimodal representations for contact-rich tasks," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 8943–8950.

[26] A. A. Shahid, L. Roveda, D. Piga, and F. Braghin, "Learning Continuous Control Actions for Robotic Grasping with Reinforcement Learning," in *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 2020-October, 2020.

[27] Z. Deng, Y. Jonetzko, L. Zhang, and J. Zhang, "Grasping Force Control of Multi-Fingered Robotic Hands through Tactile Sensing for Object Stabilization," *Sensors*, vol. 20, no. 4, p. 1050, feb 2020.

[28] J. Romero, T. Feix, H. Kjellström, and D. Kragic, "Spatio-temporal modeling of grasping actions," *IEEE/RSJ 2010 International Conference on Intelligent Robots and Systems, IROS 2010 - Conference Proceedings*, pp. 2103–2108, 2010.

[29] B. S. Zapata-Impata, P. Gil, and F. Torres, "Learning Spatio temporal tactile features with a convLSTM for the direction of slip detection," *Sensors (Switzerland)*, vol. 19, no. 3, feb 2019.

[30] H. Ravichandar, A. S. Polydoros, S. Chernova, and A. Billard, "Recent advances in robot learning from demonstration," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, 2020.

[31] Y. Lin, S. Ren, M. Clevenger, and Y. Sun, "Learning grasping force from demonstration," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 1526–1531, 2012.

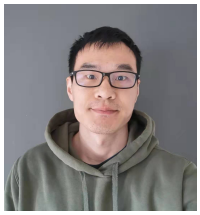
[32] E. Misimi, A. Olofsson, A. Eilertsen, E. R. Øye, and J. R. Mathiassen, "Robotic Handling of Compliant Food Objects by Robust Learning from Demonstration," in *IEEE International Conference on Intelligent Robots and Systems*, 2018.

[33] S. Ekvall and D. Kragić, "Grasp recognition for programming by demonstration," *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2005, pp. 748–753, 2005.

[34] E. De Coninck, T. Verbelen, P. Van Molle, P. Simoens, and B. Dhoedt, "Learning robots to grasp by demonstration," *Robotics and Autonomous Systems*, vol. 127, 2020.

[35] E. Forgy, "Cluster analysis of multivariate data : efficiency versus interpretability of classifications," *Biometrics*, vol. 21, pp. 768–769, 1965.

[36] J. C. Dunn†, "Well-separated clusters and optimal fuzzy partitions," *Journal of Cybernetics*, vol. 4, no. 1, pp. 95–104, 1974.



Shuaijun Wang received the M.S. degree in Mechatronics Engineering from Harbin Institute of Technology in 2015. He is currently working toward his Ph.D. degree in the state key laboratory of robotics and system, Harbin Institute of Technology (HIT), Harbin, China, and also is a visiting researcher in the Advanced Intelligent Robotics (AIR) Lab, School of Informatics, University of Edinburgh, UK. His research interests include robotic grasping, manipulation and machine learning.

Wenbin Hu received the B.S. degree in Automation from Tsinghua University, Beijing, China, in 2018. He is currently a Ph.D. student in robotics with the Advanced Intelligent Robotics (AIR) Lab, University of Edinburgh, UK. His research interests include learning based planning and control of robotic grasping and manipulation.

Lining Sun received his Ph.D. degree in Engineering from the Mechanical Engineering Department of Harbin Institute of Technology in 1993. He is a National Outstanding Youth Fund Winner, Changjiang Scholar Distinguished Professor by the Ministry of Education. He formed an internationally competitive robotics team and made outstanding contributions to the creation and development of robot-related disciplines in China.

Xin Wang received Ph.D. degree from Harbin Institute of Technology in 2013 and was a postdoctoral researcher in the Department of Advanced Robotics, Italian Institute of Technology (IIT) from 2014 to 2016. He is currently a Researcher with the Shenzhen Academy of Aerospace Technology.

Zhibin (Alex) Li is an assistant professor at the School of Informatics, University of Edinburgh. He obtained an joint PhD degree in Robotics at the Italian Institute of Technology (IIT) and University of Genova in 2012. His research interests are in creating intelligent behaviors of dynamical systems with human comparable abilities to move and manipulate by inventing new control, optimization and deep learning technologies.